ABSTRACT OF THE DISCLOSURE

The invention provides an AV signal processing apparatus and method by which a boundary between scenes is detected so that recorded video data can be played back beginning with an arbitrary scene. First, video data inputted is divided into video segments or audio segments or, if possible, into both of video and audio segments. Then, feature amounts representative of features of the segment are calculated, and then, similarity measurement between segments is performed using the feature amounts. Thereafter, it is discriminated whether or not the segment corresponds to a break of a scene. Thus, the video-audio processing apparatus uses the dissimilarity measurement criterion and the feature amounts calculated as above to determine, regarding each segment as the reference segment at present, in which one of the past and the future with respect to the reference segment the ratio of presence of neighboring similar segments is higher, and investigates the pattern of the change of the ratio to discriminate whether or not the reference segment corresponds to a boundary of a scene.